LIBRARY OF THE
UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

510.84
Il6r
no. 22-59
incompl.
cop. 2

Digitized by the Internet Archive
in 2013

http://archive.org/details/upperboundfornum58mull

UNIVERSITY OF ILLINOIS

GRADUATE COLLEGE

DIGITAL COMPUTER LABORATORY

INTERNAL REPORT NO. 58

AN UPPER BOUND FOR THE NUMBER OF CERTAIN ERROR CORRECTING CODES

BY

D. E. MULLER

August 9, 1954

An Upper Bound for the Number of Certain Error Correcting Codes

Let S be the set of n dimensional vectors whose components may take only values 0 or 1. An error correcting code is a subset T of S having the property that any pair of vectors in T differ in at least d components. The general question of how numerous a set T may be found for given n and d is of importance in information theory and will be treated in this discussion.

Fano[1] has shown that if a set T can be found having the property that every member of S lies a distance of less than d/2 from some member of set T, then the number of members of the set T approaches 2^{nK} as n approaches infinity. The quantity K is given by the formula

$$K = 1 + (\log_2 d/2n)\, d/2n + (\log_2(1 - d/2n))(1 - d/2n) \qquad (1)$$

This number 2^{nK} represents an upper bound for the number of members of the set T, and if it could be achieved the resulting set T would be an error correcting code of maximum efficiency.

Fano's formula has not been achieved by actual codes in which d/2n is held constant as $n \to \infty$, and there may even be some doubt as to whether any non-zero constant K may be found to satisfy the formula 2^{nK}. In the special case of $n = 2^p$ where p is an integer it is possible to construct a set T whenever d is a power of 2, say 2^m. The number of members of this set T is given by the formula

$$2^{\sum_{i=0}^{p-m} \binom{p}{i}} \qquad (2)$$

The sum of binomial coefficients in the exponent will be a polynomial in p of degree p-m and consequently cannot be proportional to $n = 2^p$.

(1) Robert M. Fano "Notes for Statistical Theory of Information" (1953) Page IV-34, Mass. Inst. of Technology

In the case m = p-1 we get d = 2^{p-1} = n/2 and the number of members of set T is merely 2^{1+p} = 2n. The theorem given here treats the case of general n with d = n/2 and shows that the set T can contain no more than 2n members and that the formula of Fano gives, therefore, a much higher than necessary upper bound in this case.

Let us designate the number of components in which two vectors a and b differ by the symbol L(a,b). The condition that d = n/2 may then be written as $L(a,b) \geq n/2$ if a and b are in T.

THEOREM: It is not possible to pick a subset T of more than 2n members from the set S of n dimensional vectors whose components are 0 or 1 if one requires that any pair a, b from set T satisfy $L(a,b) \geq n/2$. If the condition is made $L(a,b) > n/2$ it is never possible to pick more than n+1.

PROOF: Let $(a_0, a_1, \ldots, a_{n-1})$ represent the n component vector "a" whose components a_i are either 0 or 1. Let us adopt the notation $(A_0, A_1, \ldots, A_{n-1})$ to represent the transformed vector A, where the vector A will be transformed by the formula

$$A_i = (1 - 2a_i)/\sqrt{n} \qquad (1)$$

Thus if $a_i = 0$, $A_i = 1/\sqrt{n}$ and if $a_i = 1$, $A_i = -1/\sqrt{n}$. These are the only two possible choices for A_i and consequently

$$\sum_{i=0}^{n-1} (A_i)^2 = 1. \qquad (2)$$

We shall let V represent the broader class of vectors satisfying (2) (these are unit vectors). Transforms from set S are in V but, of course, V contains vectors which are not transforms from S. The theorem will be proven for set V and as a result will apply to the smaller set S.

Formula (1) allows a numerical representation of the quantity L(a,b). If A_i is the transform of a_i and B_i is the transform of b_i

then $1/2 - n/2\ A_i\ B_i$ will be 0 if a_i and b_i agree and 1 if a_i and b_i disagree. By summing such expressions we obtain

$$L(a,b) = (n/2)\ (1 - \sum_{i=0}^{n-1} A_i\ B_i). \tag{3}$$

This formula permits the condition $L(a,b) \geq n/2$ to be written $\sum_{i=0}^{n-1} A_i\ B_i \leq 0$ and the condition $L(a,b) > n/2$ to be written $\sum_{i=0}^{n-1} A_i B_i < 0$.

Let us assume that a set T has been found containing e members $r^0, r^1, \ldots, r^{e-1}$ which satisfy the condition $L(r^j, r^k) \geq n/2$ if $j \neq k$. The transformed set V contains e members $R^0, R^1, \ldots, R^{e-1}$ and satisfies the condition

$$\sum_{i=0}^{n-1} R_i^j\ R_i^k \leq 0 \tag{4}$$

The set of vectors V may be rotated if each vector is post multiplied by a unitary matrix M. Under such a rotation the scalar product $\sum_{i=0}^{a-1} A_i B_i$ is invariant and condition (4) will therefore continue to hold.

To show that $\sum_{i=0}^{n-1} A_i B_i$ is invariant under a rotation we need merely write $\sum_{i=0}^{n-1} A_i B_i = AB^T$. But $A' = AM$ and $B' = BM$. Hence $A'(B')^T = AM\ (BM)^T = AMM^T B^T$. M, however, is unitary so that $M^T = M^{-1}$ and $AMM^{-1}B^T = AIB^T = AB^T$.

We now show that it is possible to rotate V into V' so that $(R^0)' = (1,0,0,0,\ldots,0)$. In other words we shall perform a rotation M so that R^0 transforms into $(1,0,0,\ldots,0)$. We begin by performing a rotation by M_1, where M_1 is a unitary matrix defined as follows,

$$M_1 = \begin{bmatrix} \dfrac{R_0^0}{\sqrt{(R_0^0)^2 + (R_1^0)^2}} & \dfrac{-R_1^0}{\sqrt{(R_0^0)^2 + (R_1^0)^2}} & 0 & \cdots & 0 \\ \dfrac{R_1^0}{\sqrt{(R_0^0)^2 + (R_1^0)^2}} & \dfrac{R_0^0}{\sqrt{(R_0^0)^2 + (R_1^0)^2}} & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 \quad 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

When one post multiplies R^0 by M_1 no components are changed except R_0^0 which becomes $\sqrt{(R_0^0)^2 + (R_1^0)^2}$ and R_1^0 which becomes 0. Similarly a matrix M_2 may be constructed which alters only R_0^0 and R_2^0, making R_2^0 equal to 0 etc. The matrix $M = M_1\ M_2\ M_3,\ldots,M_{n-1}$ will thus rotate R^0 into $(1,0,0,\ldots,0)$.

We now remove $(R^0)'$ from V' and place it in a set W. If the vector $(-1,0,0,\ldots,0)$ is in V' we also place it in W, removing it from V'. We see that all remaining vectors in V' must have non-positive first components since $\sum_{i=0}^{n-1} (R_i^0)'\ (R_i^j)' = (R_0^j)' \leq 0$. Thus we may write the condition

$$\sum_{i=0}^{n-1} (R_i^j)'\ (R_i^k)' \leq 0$$

whenever $(R^j)'$ and $(R^k)'$ are in V' as

$$(R_0^j)'\ (R_0^k)' + \sum_{i=1}^{n-1} (R_i^j)'\ (R_i^k)' \leq 0\ .$$

But since $(R_0^j)'$ and $(R_0^k)'$ are both non-positive we have $(R_0^j)'\ (R_0^k)' \geq 0$

and hence $\sum_{i=1}^{n-1} (R_i^j)'\ (R_i^k)' \leq 0$ (5)

Condition (5) may now replace condition (4) with one less dimension and with one or two less vectors in V' than we had in V. It is true that the normalization condition (2) may not hold for the n-1 dimensions but condition (5) will not be affected if we multiply any vectors in V' by suitable constants to make it hold. We see that any vector in V' whose last n-1 components are zero has been removed from V' and placed in W so that normalization is always possible.

The entire process described here may be repeated for the set V' in n-1 dimensions and a new set V'' constructed. Each time this process is repeated we put at most two vectors in W and lose one dimension. Thus, after n repetitions we have placed all e vectors in W and hence e may be not greater than 2n. This proves the first half of the theorem.

To prove the second part of the theorem we follow the same process but use the condition

$$\sum_{i=0}^{n-1} R_i^j\ R_i^k < 0. \qquad (6)$$

Using this condition we see that if both the vectors (1, 0, 0,...,0) and (-1,0,0,...,0) are placed in W then no other vector may be left in V and satisfy (6). Thus only one vector is placed in W at each step prior to the last one since to place two vectors in W would cut short the process (make e less than necessary). At the last step two vectors may be placed in W since no more vectors will be left in V. Hence e may not be greater than n+1 in this case and the second half of the theorem is proved.

BOUND BY
HNM

UNIVERSITY OF ILLINOIS-URBANA
3 0112 048371618